Architecture for a Distributed, Sharable Information Security Threat Management

D.S. Bhilare
School of Computer science & IT D.A. University, Indore India

Abstract

Universities and colleges are rapidly getting dependent on their Campus Networks for their routine functioning. These networks are almost identical in their nature of operation, users, security threats being faced and potential vulnerabilities. Most of the educational institutes do not have adequate resources to identify and respond to these threats. Therefore, a distributed & sharable approach is required where information related to these incidents, attack pattern and actions taken are shared. The key benefit of a distributed collaborative approach is a global view of the malicious activities. Supplementing the information gathered locally with relevant information gathered across the globe can provide a more precise model of an attacker's behavior and activity pattern.

It is proposed that every educational institute maintains its own incident database using a common protocol, where security incidents, its analysis and solutions are recorded. The proposed architecture allows institutes of higher education to collaborate and cooperate with each other on security issues in key areas including incident response, attack mitigation, and preventive measures. This would also avoid duplication of efforts and an early solution can be found, based on past experience.

The proposed architecture improves response time in handling the information security incidents. In addition, it decreases the risk of attack through collaborative projects, customized to meet each participant's Information Security needs. It is also proposed that based on continuous analysis of available data by experts, warning alerts or preventive measures will be communicated to the member institutions proactively to mitigate existing or future threats.

KEYWORDS : Component; Information security; collaborative security; OAI-PMH; harvested approach

INTRODUCTION

Information Security Incident Management has become an important component of overall information Security Management System. New types of security related incidents emerge every day. Preventative activities based on the results of risk assessments can lower the number of incidents, but not all incidents can be prevented. An incident response capability is therefore necessary for rapidly detecting incidents, minimizing loss and destruction, mitigating the weaknesses that were exploited, and quickly restoring the IT services.

Because performing incident response effectively is a complex process, establishing a successful incident response capability requires substantial planning and resources. Establishing clear procedures for assessing the current and potential business impact of incidents is critical. Equally important is implementing effective methods of collecting, analyzing, and reporting data. Building relationships and establishing suitable means of communication with other internal groups (e.g., human resources, legal) and with external groups (e.g., other incident response teams, law enforcement) are also vital [20].
The proposed architecture considers these issues while designing the solution, detail explanation of the architecture is given in the following sections.

RELATED WORK

Georgia Tech Information Security Center (GTISC), one of the leading academic research centers focused on information security, believes strongly that a proactive and collaborative approach to understanding emerging threats will help us develop more effective information security technologies and strategies [16].
Earlier research on distributed threat management has focused mainly on the exchange of data within a single organization. Focus was on distributed collection and centralized correlation [11]. The DShield project [12] is an example of a centralized repository that receives intrusion alerts from many distributed sources. Cuppens and Miege [13], [14] discuss methods for cooperatively correlating alerts from different types of intrusion detection systems.
A study showing resources needed to counter worm propagation, concludes that a response needs to be mounted in 2-3 minutes and that participation of nearly all nodes required to be effective [15]. They also suggest that the Internet by its inherent weaknesses is not effective in preventing and containing worm outbreaks. Their results suggest that both technological and administrative issues must be addressed before any effective defense can be mounted against such Internet-wide threats. While these numbers are specifically derived in the context of quarantining Internet worms, their results show that current and foreseeable threats demand a cooperative and collaborative approach to achieve desired level of security.
In view of the current growth trends in the Campus Wide Networks particularly in the institutes of higher education where most of the universities & colleges have already built or in the process of building there own digital libraries and information assets; there is a need for a mechanism which can protect sensitive information from possible threats. This involves recording security incidents; match it with existing patterns available in the cooperative incident database which is constantly being updated by various participating educational institutes. There should be a single-window system from users point of view i.e. a query submitted should fetch and present the information from all the incident databases who are member of the cooperative.
In order to achieve the above mentioned objectives a common protocol is needed which enables access to Web-accessible material through interoperable repositories for metadata sharing, publishing and archiving. Here, we have identified OAI-PMH protocol, which can fulfill our requirements. It arose out of the Open Archives Initiative (OAI), the OAI

develops and promotes a low-barrier interoperability framework and associated standards, to access digital materials. As it says in the OAI mission statement "The Open Archives Initiative develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content."

The OAI-Protocol for Metadata Harvesting (OAI-PMH) defines a mechanism for harvesting records containing metadata from repositories. The OAI-PMH gives a simple technical option for data providers (participating universities) to make their metadata available to services, based on the open standards HTTP (Hypertext Transport Protocol) and XML (Extensible Markup Language). The metadata that is harvested may be in any format that is agreed by a community, although unqualified Dublin Core is specified to provide a basic level of interoperability. Thus, metadata from many sources can be gathered together in one database, and services can be provided based on this centrally harvested, or "aggregated" data. The link between this metadata and the related content is not defined by the OAI protocol. It is important to realise that OAI-PMH does not provide a search across this data, it simply makes it possible to bring the data together in one place. In order to provide services, the harvesting approach must be combined with other mechanisms.[1]

A browse interface is very difficult to build when the metadata to be browsed is distributed across a number of repositories. Therefore, preferred approach would be to get all the metadata records together in one place.

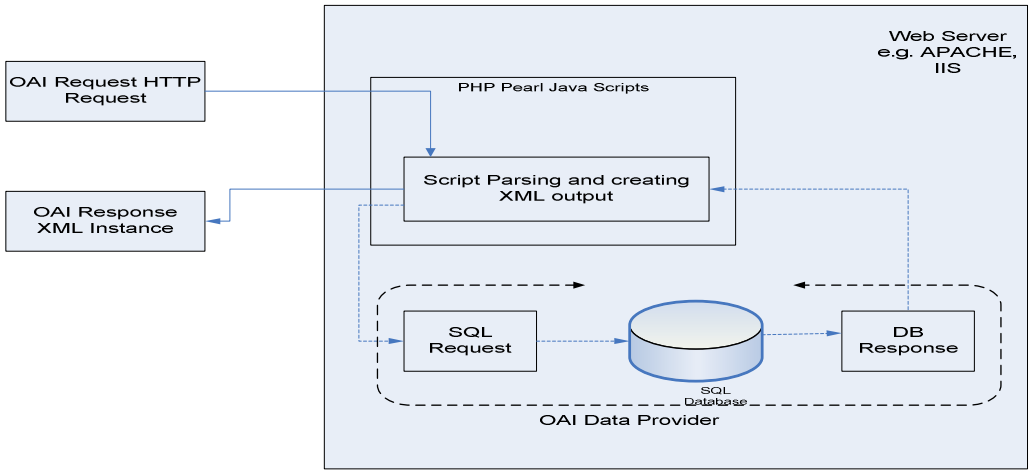

Fig. 1: System Architecture of OAI-PMH

ACCESS METHODS

There are broadly two main approaches for accessing the global collaborative distributed database to get the required information: Federated Approach and Harvested Approach.
The major issues are as follows:

Several issues must be addressed prior to implementation of Distributed Collaborative

Inter University Threat Management Mechanism [18].

- There should be a common protocol allowing sharing of information which is at present stored in a different formats.
- There should be a central service provider who can receive requests, interact with participating organizations and return the best result to the client.
- Exchanging alert data in a full mesh network increases bandwidth requirements.
- Information exchange needs to be carefully managed. Confidential information should not leak.
- Finding the most appropriate databases from which the user request can be successfully satisfied and Selection of the most appropriate solution in the given situation from the available alternatives
- Integration of all the results into an agreed upon global format that can be consistently presented to the end user.
- Handling duplication of results

Digital library experience suggested that cross searching does not scale well, at least partly because the search service degrades to the level of the slowest and least reliable server in the cross search set [19]. For example, NCSTRL found that distributed searching of a small number of nodes was viable, but that performance was very bad over 100 nodes. The more servers are cross-searched, the higher are the chances of encountering one or more slow or unreliable servers.

Our proposed protocol uses harvested access to establish collaborative incident information system. In a harvested access, generally called harvesting, all the metadata from all the databases present at distinct remote places is gathered at a common location. This metadata information then plays a vital role to find out the required information requested by the user.

With harvesting, we have less network traffic as the metadata from all the databases is presented well in advance to the central common place and thus depending on the fired query, it can be decided which database to be searched for finding the results. Moreover, harvesting is a robust light weight protocol.

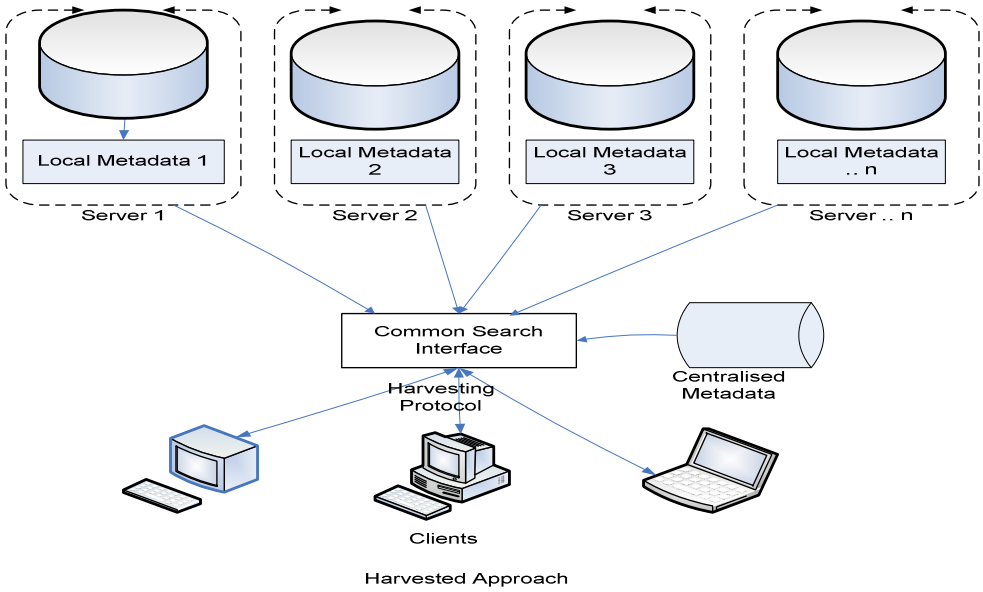

Fig. 2: Functioning of Harvested Approach

On the basis of the above discussion, harvested approach is preferred over the federated approach in the proposed architecture. The following section gives the detailed blue print of the proposed architecture for distributed collaborative inter university incident handling mechanism.

PROPOSED ARCHITECTURE

The proposed architecture for implementing a distributed, sharable Information Security Threat Management is built around open source OAI-PMH protocol. The overall structure is divided into four key modules: Incident Data Capture Module, Metadata Generator Module and Service Provider Module. As discussed earlier, harvested approach is preferred over federated approach for integration and dissemination of metadata. For a harvested approach it is required that the incident databases should expose the metadata related to the stored information as per the metadata standard used by the central interface. Here it is important that no other information should be accessible to the users.

Design of the Incident Database

- The proposed incident database stores the following information about the each incidence:
- Incident ID.
- University/College ID
- Who attacked ?
- Why was the attack performed?

- When did it happen?
- How did they do the attack?
- How widespread is this incident?
- Did this happen because of poor security practices?
- What steps are being taken to determine what happened and to prevent future occurrences?
- What is the impact of this incident?
- Was any personally identifiable information exposed?
- What is the estimated cost of this incident?

Components of the Architecture

The complete architecture is illustrated in the Fig. 3. There are four key modules as shown below:

- Incident Data Capture Module (IDCM)
- Metadata Generator Module (MGM)
- Service Provide Module (SPM)
- Analyzer Module (AM)

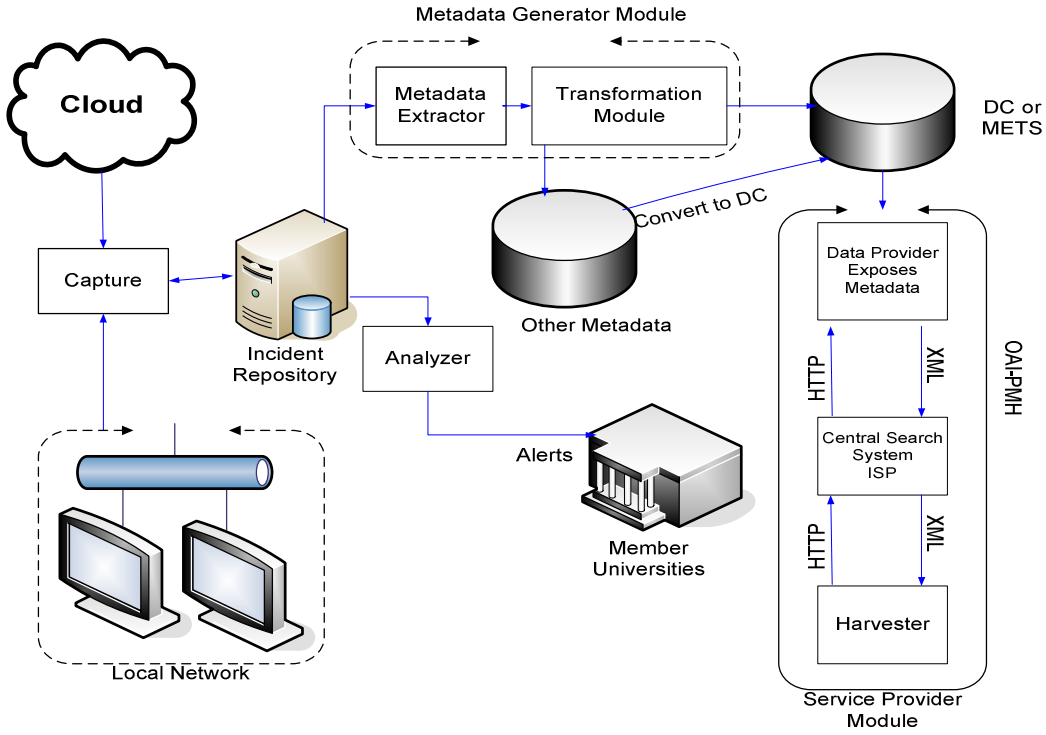

Fig. 3: An Architecture for Distributed, Sharable

Incident Data Capture Module (IDCM)

The "Capture" module interacts with the environment, which includes the External World and Campus Network. Whenever any new input or suspicious activities is detected, it is recorded in the incident database in the pre-decided format, if it is not already recorded. Under certain circumstances, the incident database is also updated manually, as sometimes the "capture" module is unable to detect all the incidents.

The "capture" module also maintains a log of all the incidents so that it can be used by the administrator to formulate the policies for the incident management and reporting.

Metadata Generator Module (MGM)

The Metadata Generator Module (MGM) module continuously gathers input from the incident database and generates corresponding metadata. The MGM, manages the mapping relationships between the information present in the incident database to the standards like DC (Doublin Core) or METS (Metadata Encoding and Transmission Standard).

This metadata is generated depending on the contents of the incident database and other input. This is carried out by checking how frequently a term is used in the incident reporting text. The mapping is thus completely dependant on the contents of the text.

The metadata thus generated by the Extractor sub module forms raw input to the Transformation sub module. The Transformation sub module separates metadata formats recognized by the OAI-PMH in a separate repository, which mainly contains DC and METS. The remaining formats are transferred to other formats, which later on converted to accepted format using available metadata conversion tools. Finally, these metadata formats which are essentially in XML format are made available to the Data Provider.

Service Provider Module (SPM)

Whenever a member university or college requires some information from the member community, a request is fired by local client having software called harvester. The harvester software is required to issue the OAI-PMH request for fetching the metadata of incident databases. This request is first sent to the central service provider. The Service Provide tries to resolve the request using local repository. If required information is not available locally, then it is passed to the member repositories. There is a module called "data provider" at the local database end which exposes this metadata to the harvester via the same service provider.

Hence in OAI-PMH, it is the data provider who does the job of exposing the metadata to the harvester. Any institute that wants to be a member of the collaborative incident databases group has to first ensure that they follow the same

protocols to expose their metadata information to the central service provider.

Analyzer Module (AM)

The Analyzer module based on standard vulnerability database and history analysis sends alerts to member universities. Analyzer can be configured manually also to send emergency global messages to the member institutes.

Thus the "distributed collaborative inter university incident management system" can be built by following the proposed architecture. We have demonstrated how metadata can be generated and send to the data provider for exposing it to the service provider.

Post Incidence Activities for Future Planning and Management

Using Collected Incident Data History for Future Planning and Management

Lessons learned activities should produce a set of objective and subjective data regarding each incident. Over time, the collected incident data should be useful in several capacities. The data, particularly the total hours of involvement and the cost, may be used to justify additional funding of the incident response team. A study of incident characteristics may indicate systemic security weaknesses and threats, as well as changes in incident trends. This data can be put back into the risk assessment process, ultimately leading to the selection and implementation of additional controls. Another good use of the data is measuring the success of the incident response team. If incident data is collected and stored properly, it should provide several measures of the success (or at least the activities) of the incident response team. Furthermore, organizations that are required to report incident information will need to collect the necessary data to meet their requirements.

IMPLEMENTATION

Implementation Strategy for OAI-PMH Grammar

In order to access OAI-PMH repositories for harvesting metadata use of the Pearl based toolkit, Net::OAI::Harvester[17] is proposed. Implementation code for one verb of OAI-PMH is given below. Similarly, remaining verbs can also be implemented:

The OAI-PMH is essentially a set of request/response messages which may be sent over HTTP to retrieve metadata that is encoded in XML. So one can construct a familiar URL and get back an XML document containing the required metadata. From a programming perspective there are several issues that arise when writing a OAI-PMH harvesting program: HTTP requests need to be URL-encoded for safe transmission; error conditions can arise which must be handled gracefully; resumption tokens may be used to break up a response into chunks. Of greatest concern here is that all responses are arbitrarily large XML documents.

Net::OAI::Harvester is a Perl module that abstracts away all the details of generating

the HTTP request, handling error conditions, and parsing XML so that extracted data can be easily used.

After installing Perl, Net::OAI::Harvester can be installed with one command:

```
perl -MCPAN -e 'install Net::OAI::Harvester'
```

Implementation of "Identify" verb

Using the following script, identification of the repository can be obtained:

```
1 use Net::OAI::Harvester;
2 my $var0 = Net::OAI::Harvester->new(
3 baseURL => 'http://archive.upper.net/cgi-bin/mph1' );
4  my $identity = $var0->identify();
5 print $identity->repositoryName(),"\n\n";
```

OUTPUT:

Central Library Vikram University OAI-PMH Repository

Code for capturing metadata

The OAI-PMH has three verbs which facilitate obtaining metadata from a repository: ListIdentifiers, ListRecords and GetRecord. Each of these verbs translates into a Net::OAI::Harvester method: listIdentifiers(), listRecords() and getRecord(). The OAI-PMH defines an identifier as unambiguously identifying an item within a repository. The idea of the ListIdentifiers verb is that it allows a harvester to see what identifiers exist in the repository and to only request those that are of interest. The following code extracts metadata related to physics department stored in DC (Doublin Core) format. The "while" loop extracts all the metadata records, reading block of records at a time till end is reached, for the specified set i.e. physics.

```
1  use Net::OAI::Harvester;
2   my $var1 = Net::OAI::Harvester->new(
3baseURL   => 'http://dauniv.ac.in/meta1');
4  my $list = $var1->listIdentifiers(
5  metadataPrefix  => 'meta_dc',
6      set   => 'phy');
7      while ( my $header = $list->next() ) {
8            print $header->identifier(),"\n"; }
```

OUTPUT:

```
oai:arXiv.org:cmp-lg/9404001
oai:arXiv.org:cmp-lg/9404002
oai:arXiv.org:cmp-lg/9404003
    ...
```

Thus we can implement the proposed architecture using open source protocols and tools.

CONCLUSION AND FUTURE WORK

The proposed architecture improves institutions capability to manage Information Security threats, as more informed decisions are taken. Institutions are able to take preemptive actions without having been directly attacked. Since most of the educational institutes have limited resources to detect and respond to these threats, the proposed approach allows sharing of information related to these attacks and possible solutions. The key benefit of a collaborative approach is a better view of global network attack activity. Augmenting the local information with information gathered across the globe can provide a more precise model of an attacker's behavior and attack pattern. The proposed architecture enables collaborating educational institutes to expose the security related information of common interest, which institutes feels safe to disclose. This is implemented using an open source protocol OAI-PMH, where as the underlying database may be in different format. The participating members of these cooperative can find information about measures taken in case of similar incidents taken place in other institutes. This would avoid duplication of efforts and an early solution can be found out based on past experience.

The proposed architecture helps participating institutions in improving response time in mitigating information security incidents and emergencies. It also decreases the risk of attack through collaborative projects customized to help meet each participant's information security needs. It is also proposed that based on continuous analysis of historical data by experts, a set of best practices as preventive measures are also made available to the member universities. Further, warning alerts are communicated to the member institutions proactively for new threats or information of common interest.

One limitation of this approach is that the analysis of data captured may not reveal every possible threat, as it is very difficult to ensure that all the significant events are recorded and analyzed properly. There are incidents which may remain unnoticed for long time.

Future research may focus on data analysis towards predictive security measures and methods for low cost alert distribution and broadcast. There are issues related to bandwidth while messages are exchanged across participating institutes. Sometimes a single weak link may become bottleneck while aggregating information from collaborating institutes. Standardization need to be extended further, up to database level, which will result in better information retrieval system. Another issue would be development of a metric for ranking the threat level for particular verticals, or groups of organizations.

REFERENCES

1.The Open Archive Initiative. http://www.openarchives.org
2. Dublin Core Metadata Initiative. http://purl.org/DC/

3. Dspace digital library. http://www.dspace.org
4. L. A. Gordon, M. P. Loeb,W. Lucyshyn, and R. Richardson. CSI/FBI computer crime and security survey, 2006.
5. Howard and T. Longstaff. A common language for computer security incidents, 1998.
6. ISO/IEC. ISO/IEC 27001:2005 information technology security techniques information security management systems requirements, 2005.
7. Bhilare DS and Kawale Jai. An architecture for cooperative digital libraries. CSI Computing 2005.
8. M. Junginger, A. Balduin v., and H. Krcmar. Operational Value at Risk und Management von IT-Risiken. WISU - Das Wirtschaftsstudium, (3):356–364, 2003.
9. S. E. Schechter. Computer Security Strength & Risk: A Quantitative Approach. PhD thesis, Harvard University, Cambridge, MA, 2004.
10. K. J. Soo Hoo. How much is enough? A risk management approach to computer security. http://iisdb.stanford.edu/pubs/11900/soohoo.pdf, June 2000.
11. S. Stolfo, “Worm and Attack Early Warning: Piercing Stealthy Reconnaissance," IEEE Privacy and Security, May/June 2004.
12. J. Ullrich, “Dshield home page." http://www.dshield.org/, 2004.
13. F. Cuppens and A. Miege, “Alert Correlation in a Cooperative Intrusion Detection Framework," in IEEE Security and Privacy, 2002.
14. F. Cuppens and R. Ortalo, “Lambda: A language to model a database for detection of attacks," in Proceedings of the Third International Workshop on the Recent Advances in Intrusion Detection (RAID 2000), (Toulouse, France), October 2000.
15. D. Moore, C. Shannon, G. Voelker, and S. Savage, “Internet Quarantine: Requirements for Containing Self Propagating Code," in INFOCOM 2003, 2003.
16. GTISC http://www.gtisc.gatech.edu/pdf/CyberThreatsReport2009.pdf
17. Ed Summers, "Building OAI-PMH Harvesters with Net::OAI::Harvester" ,30-January-2004 Publication: Ariadne Issue 38 http://www.ariadne.ac.uk/issue38/summers/intro.html
18. Michael E. Locasto, Janak J. Parekh, Angelos D. Keromytis, Salvatore J. Stolfo, “Towards Collaborative Security and P2P Intrusion Detection”, Proceedings of the 2005 IEEE Workshop on Information Assurance and Security T1B2 1555 United States Military Academy, West Point, NY, 15{17 June 2005, www.cs.columbia.edu/~angelos/Papers/2005/iaw.pdf
19. Open Archives Forum. (n.d.). OAI for beginners, the Open Archives Forum online tutorial. 2. History and development of OAI-PMH. Retrieved April 21, 2007, from http://www.oaforum.org/tutorial/english/page2.htm.
20. Nist Publications: computer security resource centre, SP800-61, src.nist.gov/publications/nistpubs/,